High-Density Distributed Sensing for Chemical and Biological Defense

Adam Norige, Jason Thornton, Curran Schiefelbein, and Christina Rudzinski

A new sensing architecture is being developed for plume-based chemical and biological (CB) threat detection, mapping, and prediction. This effort seeks to provide a novel, affordable, and effective solution to spatially distributed CB detection by working in four key areas. These four areas include the analysis of performance gains offered by distributed sensing; the development of detection algorithms; the construction of an inexpensive sensor prototype; and experimentation with inexpensive, low-power deployment infrastructures. This multipronged approach to inexpensive spatially distributed sensing has led to the creation of a robust CB sensing architecture that offers many significant advantages over traditional approaches.

As the nation faces a growing concern regarding asymmetrical threats, such as terrorist attacks using chemical and biological (CB) weapons, the diversity of environments requiring protection is on the rise. CB sensor systems, once reserved for military battlefield deployments, are now appearing in civilian structures such as public transportation systems and office buildings [1]. This expanding range of sensing environments and operating conditions is forcing the requirements for protective sensor systems to evolve. Not only do future detection systems have to satisfy traditional requirements such as sensitivity, response time, probability of detection, and false-alarm rates, but they must also satisfy other constraining factors such as cost, power consumption, and maintainability [2]. Ultimately, operators seek low-cost detection systems with flexible deployment capabilities that do not sacrifice overall detection performance.

To address these evolving CB detection requirements, Lincoln Laboratory is developing a CB sensing framework, consisting of high-density networks of inexpensive point sensors. Not only does this sensing framework address the evolving detection requirements, but it also has the potential to provide many advanced capabilities, such as threat tracking, mapping, and prediction, which are not readily feasible with many traditional frameworks. Throughout our development efforts, we viewed high-density sensor networks as an interdependent system of sensors, network detection algorithms, and deployment infrastructure technologies. Viewing this new framework as a system of technologies is critical because the key technical challenges exist within the system's inter-

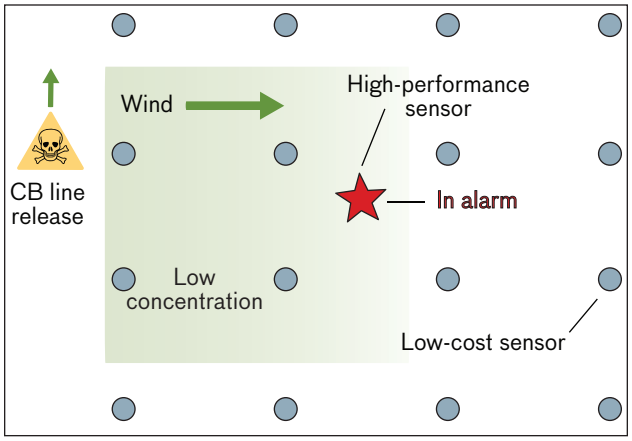


FIGURE 1. Line releases usually consist of low concentrations of material distributed across a wide area, thus making sparse deployments of high-fidelity sensors better suited for detection.

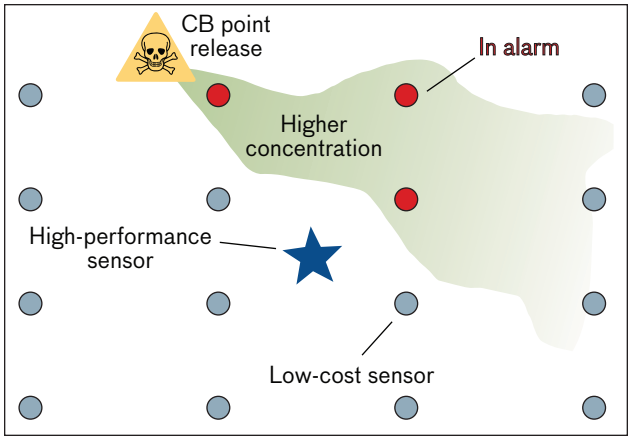


FIGURE 2. In a point-release scenario, high-density deployments of low-cost sensors have a higher probability of encountering the chemical-biological (CB) plume than the single high-fidelity sensor.

dependencies. For example, as more sensors are added to the network, the need arises to establish agile deployment mechanisms such as wireless communication links. Dealing with wireless communications, especially in the low-power domain, introduces bandwidth constraints on data transfers within the network, thereby restricting the amount of data available to the network detection algorithms and ultimately influencing the overall detection performance. To maintain a systems-based approach, our development efforts focused on four key areas: the performance analysis of high-density sensor networks, the development of inexpensive sensor technologies, the development of network detection algorithms, and the use of agile deployment technologies.

Chemical and Biological Threats

Fundamentally, the majority of large-scale outdoor CB attacks fall within two categories, line and point. These two categories refer to the style of release. Line releases, as shown in Figure 1, consist of a moving release point that disseminates a low concentration of threat material over a wide area. Sensing CB threats from line releases requires a few high-fidelity sensors capable of detecting very small quantities of threat material. Sparse deployments of high-fidelity sensors are suitable for detecting line releases because low concentrations of threat material permeate a wide area and have a high probability of encountering a sensor.

Conversely, a point release, shown in Figure 2, consists of a stationary release point that disseminates a high concentration of threat material either as a burst or as a continuous release. Point releases generally target specific areas, but can affect large areas, depending on the atmospheric conditions. Unlike with line releases, sparse sensor deployments are not well suited for detecting CB threats from point releases because individual CB threat plumes from point sources consist of distinct unmixed regions of both high and low concentrations [3]. Individual plumes do not have Gaussian concentration distributions or smooth gradients between regions of high and low concentrations, which are seen in the average of many plumes or in plume ensembles. As shown in Figure 3, such plumes can meander and exhibit unpredictable concentration distributions, making single-sensor placement for threat detection a difficult task. Even an exquisitely sensitive sensor will never detect threats that do not interact with it.

On the basis of the understanding of the behavior of individual plumes, high-density sensor networks are an ideal choice for detecting CB threats from point releases. High-density sensor networks work well because the increased sensor density offers increased sensor coverage within the spatial domain, thus increasing the probability that the plume will interact with at least a few sensors.

Because, in most cases, predicting whether an adversary may use a point or line style of attack is impossible, it may make sense to always use high-density networks of high-fidelity sensors. In theory, this approach is ideal for addressing the entire threat space, but is greatly hindered by practical considerations. Currently, the high-fidelity

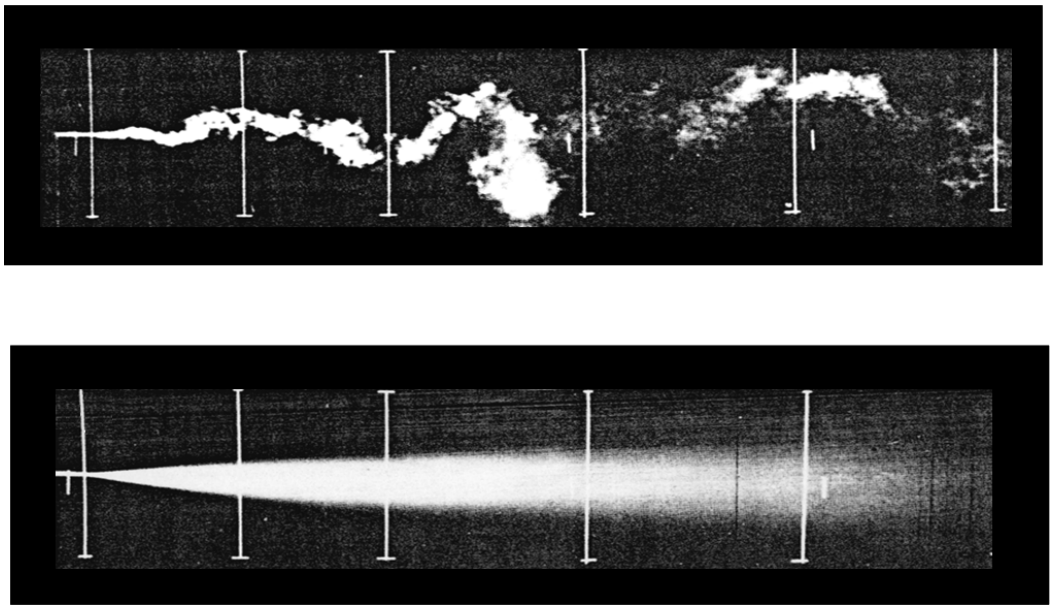


FIGURE 3. A photograph comparison reveals the difference between an individual plume (top) and an averaged set of plumes (bottom) [3]. Individual plumes do not have Gaussian concentration distributions or smooth gradients between regions of high and low concentrations, which are seen in the average of many plumes or in plume ensembles.

sensors that are capable of detecting ultralow concentrations of threat material can cost upwards of $50K per sensor. Building high-density networks of hundreds of sensors is immediately prohibited by the involved sensor costs.

In order for high-density networks to serve as a practical solution, they must utilize inexpensive sensors. With CB sensors, there is a very strong inverse correlation between sensor cost and sensor performance. In general, as the sensor costs decline, the sensor's sensitivity diminishes and its false-alarm rate increases. Much of the recent sensor development from the CB community has focused on the development of sensors with ultralow sensitivities and false-alarm rates, but at an ever-increasing sensor cost.

It is our belief that when used intelligently in networks, inexpensive lower-performing sensors will serve as an effective solution for detecting threats from point releases. Through our research, we have determined this to be true for two fundamental reasons. First, for releases within or near the sensor network, some of the sensors are bound to be located near the threat source, which is typically a region of high concentration of material—thereby reducing the sensor's performance requirements. Second, by searching for spatially correlated sensor alarm activity within the spatial network, the network false-alarm rate is reduced by rejecting uncorrelated alarms from individual sensors. In essence, spatial networks of low-performance, inexpensive sensors can perform just as well as, if not better, than a single high-performance, expensive sensor by leveraging both the proximity of sensors to the threat source and the spatial correlations of sensor readings. Furthermore, sensor networks can be designed to be more tolerant of single-node failures and more forgiving of exact sensor placement. Much of the work discussed in this report demonstrates these concepts under realistic conditions.

Modeling and Analysis

When assessing the performance of CB sensing systems, many researchers traditionally use modeling and simulation tools, such as the Defense Threat Reduction Agency's (DTRA) Hazard Prediction and Assessment Capability tool or the Navy's Vapor, Liquid, and Solids Tracking tool [4, 5]. However, these modeling tools rely upon the Second-order Closure Integrated Puff (SCIPUFF) model to simulate the atmospheric dispersion of vapor and aerosol materials. SCIPUFF is a Lagrangian puff dispersion model that represents material dispersions as distributions of the probability of encountering a specific concentration at a specific time and location. In essence, these probability distributions represent an ensemble of plumes rather than an individual plume. The difference between plume

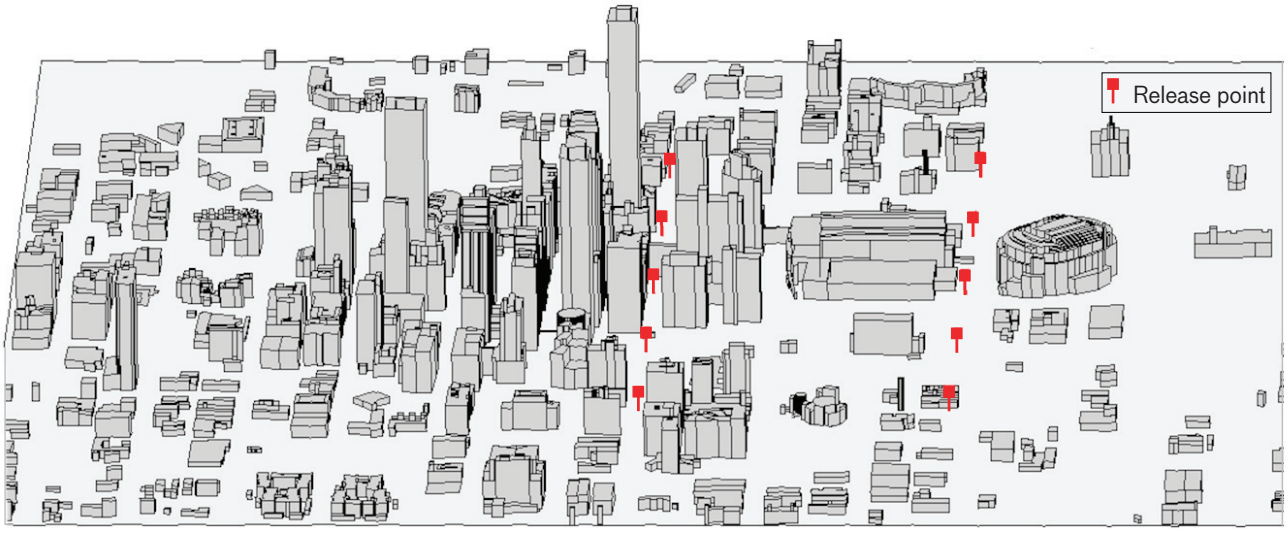


FIGURE 4. This visualization depicts the 3D virtual urban area within the computational fluid dynamics (CFD) model and the sarin release points. The urban area is derived from geographical information system data on Oklahoma City's central business district. We released 0.18, 1.8, and 18 kg each at ten different locations within the simulated environment to constitute the thirty different plumes. The staggered release times, plus the varying winds, helped increase variance among plumes.

ensembles and an individual plume is shown graphically in Figure 3. While ensemble plume models are useful for studying the full range of probable plumes under specific atmospheric and release conditions, they are not, nor were they intended to be, the best tool for studying individual plume realizations.

This is an important consideration because high-density CB sensor networks exploit the unmixed nature of individual plumes. To correctly assess the performance of high-density CB sensor networks, we needed a dispersion model capable of generating individual plume realizations. Currently, computational fluid dynamics (CFD) models are one of the most accurate methods for creating high-fidelity models of individual plumes, but are typically complex and computationally expensive and are usually built for very specific modeling tasks. These characteristics can make the development of CFD models difficult and time-consuming. Fortunately, we were able to work with Computational Fluid Dynamics Research Corporation (CFDRC) to develop a version of their "Urban Areas" model to simulate multiple instances of individual plumes [6]. CFDRC's Urban Areas framework uses CFD calculations to simulate the transport and dispersion of chemical, biological, radiological, nuclear, and explosive (CBRN/E) materials within three-dimensional urban environments. The Urban Areas simulation framework can generate plumes that closely approximate individual, realistic plumes and has undergone an extensive validation process using real field measurements. CFDRC validated the underlying models in the Urban Areas framework using data recorded from the Joint Urban 2003 (JU2003) field test, which took place in Oklahoma City in July 2003 [7]. JU2003 was sponsored by the Department of Homeland Security, DTRA, and the U.S. Department of Energy, and was designed to study the atmospheric transport of materials by conducing multiple releases of inert tracer gases within Oklahoma City's urban environment [8].

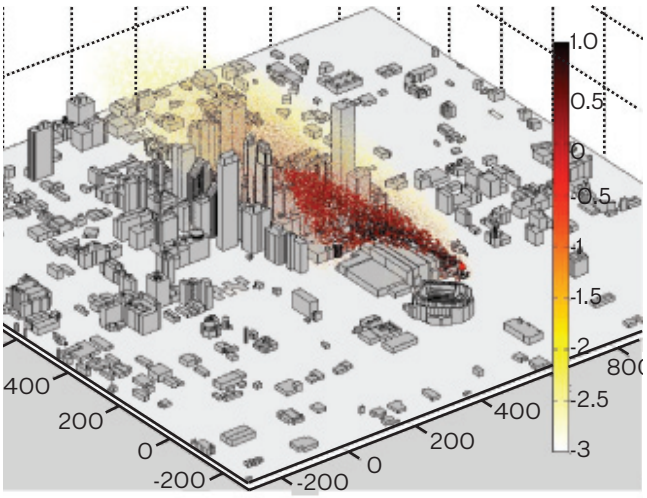


FIGURE 5. A sarin release is visualized within the CFD model. The color of the plume indicates the instantaneous sarin concentration.

By using this simulation framework, we were able to simulate thirty outdoor releases of sarin within Oklahoma City's central business district. We selected sarin as a release material because it is a chemical warfare agent that is lethal in small doses and has been used to attack urban environments in the past [9]. We extracted hypothetical sensor data from the CFD model's wind and concentration fields by using a custom-built post-processor that outputs time-series data from any coordinate within the model's virtual environment. This capability allowed us to build networks of hypothetical CB sensors and anemometers anywhere within the model. The thirty different plume dispersions along with the post-processing tools formed a highly capable test platform on which we could assess the detection performance of high-density sensor networks. Figure 4 displays the three-dimensional simulation environment and the ten unique release points. A simulated sarin release is shown in Figure 5.

Required Sensitivity versus Sensor Density. Central to the application of high-density sensor networks is the notion that as the sensor density increases, the required sensitivity for the individual sensors can be reduced while maintaining a fixed detection performance. This reduction in required sensitivity permits the use of lower-cost sensors, as illustrated in Figure 6. As more sensors are added to the network, increasing the sensor density, the sensitivity requirements for the individual sensors are significantly reduced. Note that the values on the axes in Figure 6 are specific to the modeled environment, but the relationship between sensor density and required sensitivity holds true for the general scenario of point-release detection.

Application-Specific Performance Gains. The flexibility of the CFD simulation framework allowed us to analyze specific questions regarding the application of high-density sensor networks. A pertinent application of high-density CB sensing is for fixed-site or facility protection. When it comes to this type of application, a commonly asked question is: With my fixed budget, should I deploy a small number of higher-cost, higher-performing sensors or a large number of lower-cost, lower-performing sensors? This question is not easily addressed because of the wide variety of threats that may be of concern, but for defense against point attacks, the cost-benefit analysis (summarized in Figure 7) shows that having more low-cost sensors is preferable to fewer higher-cost sensors.

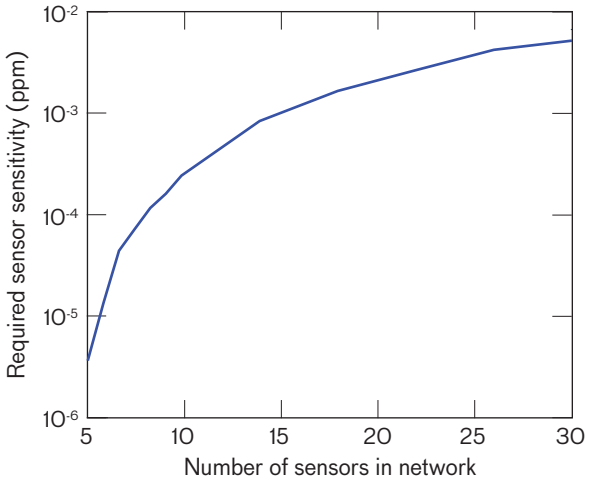


FIGURE 6. The addition of more sensors reduces the required sensitivity to detect 100% of all releases. The sensor networks were generated by randomly deploying *n* sensors across the CFD model's virtual environment at a fixed height of 3 m above ground level.

Sensor Placement. In addition to the choice between few high-fidelity sensors and many inexpensive sensors, the question of sensor placement is an important consideration when deploying sensors. If you only have a few sensors, you had better place them in the right locations; otherwise, you may miss threats completely. The most convenient approach to sensor placement is a random, or ad hoc, placement of sensors rather than an optimized sensor configuration. An optimized sensor configuration requires an in-depth analysis of all potential threats, which is expensive and time-consuming. Conversely, an ad hoc configuration can be deployed quickly and without prior knowledge of quirks of the local environment. Figure 8 shows that increasing the sensors' spatial density diminishes the required sensor sensitivity between an optimal and an ad hoc sensor configuration. The ability to rely upon an ad hoc sensor configuration greatly reduces the risk and effort involved in sensor placement.

Field Tests

When testing the performance of CB sensor systems on individual plumes, there is no substitute for field tests. Field tests enable experimentation with real plumes and ambient backgrounds influenced by real atmospheric conditions. Obtaining approval to release large quantities of materials into the atmosphere within urban environments for scientific experiment is a difficult process, so we chose

to conduct releases of tracer material in a large open field away from urban structures. In addition, we selected water-based smoke as the tracer compound because it is relatively safe and its dispersion is visually observable. Figure 9 displays our test grid and sensor nodes. We also note that out of the twenty-five sensors, some failed to report readings for at least part of the trials. Since there is usually a non-negligible risk of sensor failure in real fielded systems, distributed deployments offer the advantage of fault tolerance (in the sense that they are less susceptible to information blackout than single-node deployments).

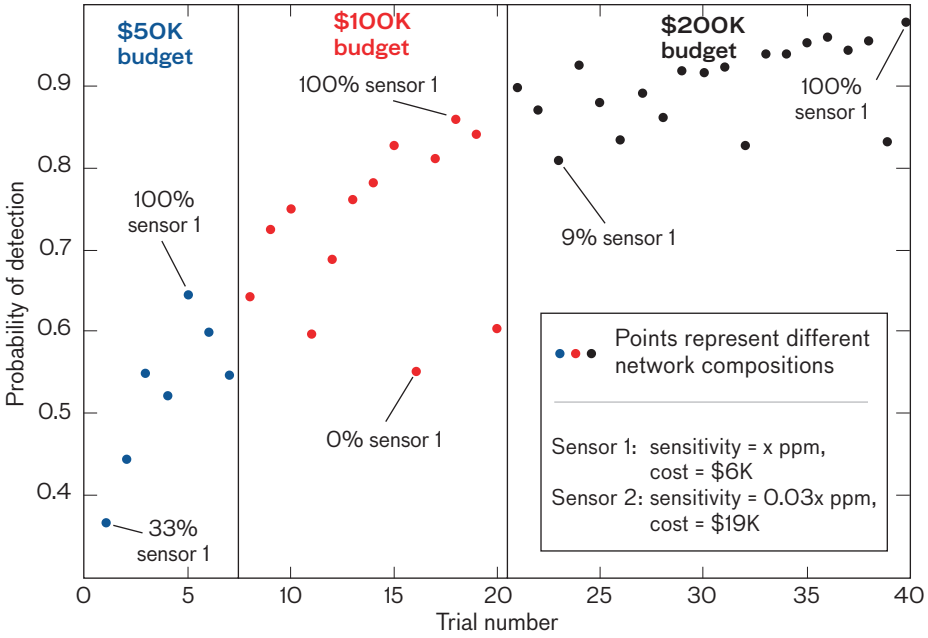


FIGURE 7. This cost benefit analysis shows the probability of detection for sensor network compositions for three different procurement budgets. Each data point represents a sensor network configuration consisting of a mixture of two commercially available chemical sensors, one high cost and one low cost with corresponding performance. For every budget, the networks consisting solely of low-cost sensors achieve the highest overall probability of detection.

Making the Field Data More Realistic. Because the smoke releases consisted of particle concentrations that were significantly higher than the particle counts observed within the ambient background, we devised a method for suppressing the signal strength of each release, so that we could experiment with more challenging concentrations while retaining fully realistic plume behavior. We controlled the signal strength by adjusting the particle counts to amplify or suppress the evidence of contaminant presence. To do this, we used ground truth about the smoke releases to estimate the number of counted particles that originated from the release and not from interferents. Then we adjusted the data by including only a specified fraction of the smoke particles in the total particle counts. Figure 10 shows a time series of particle counts at several different signal strengths for a sample node and contaminant release.

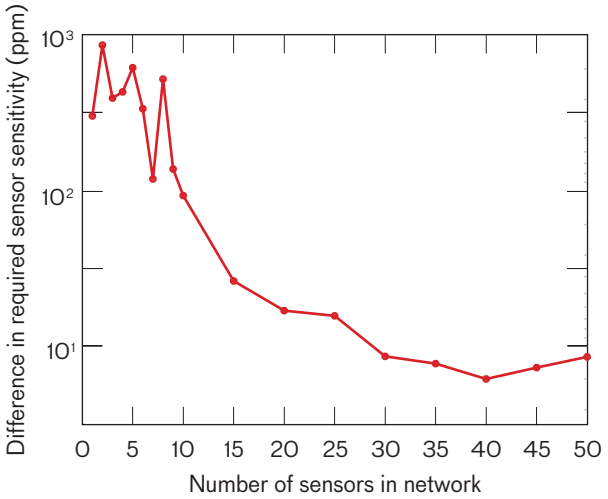


FIGURE 8. Comparison of optimal versus an ad hoc sensor placement identifies the necessary sensor sensitivity for each network configuration to achieve a probability of detection (PD) of 1. The *y*-axis is the difference between this required sensitivity for the optimal and ad hoc networks. The optimal configuration is the best possible sensor placement if one had complete knowledge of the plume's future trajectory—information that is unattainable in real scenarios. For networks consisting of few sensors, the optimized placement requires the individual sensors to be two to three orders of magnitude more sensitive than the sensors within the optimal configuration. As the sensor density increases, the required sensor sensitivity for the ad hoc configuration approaches that of the optimal sensor configuration, making an ad hoc configuration achievable.

FIGURE 9. The outdoor field tests consisted of 25 sensor nodes arranged on a square grid, spaced 50 m apart. Each node consisted of one battery-powered, tripod-mounted particle counter (either a TSI AeroTrack 8220 or MetOne HHPC-6 EX) and one anemometer (Davis 7911). Wind speed and direction measurements were transmitted in real-time from every sensor node to a base station by using low-power mesh radios (Crossbow Mica2). Sonic anemometers were placed in two locations on the grid for high-fidelity wind data to compare to the Davis 7911 anemometer readings. During the test, 36 smoke candles that produced 40,000 ft^3 of water-based smoke in 30 s were activated. In addition to the concentrations and spread of the 36 smoke releases, measurements of the ambient background were collected periodically throughout the test.

We evaluated detection performance in terms of both successful detections and false alarms. We declared a successful detection of a contaminant release if the contaminant score exceeded an alarm threshold at least once during that release. Correspondingly, we declared a false alarm every time the contaminant score exceeded the same threshold during periods of background data collection. Note that either event depends upon the alarm threshold. In the ideal case, there exists a threshold that yields successful detection of all releases while triggering no false alarms. In fact, we can achieve this ideal performance on the original, unadjusted data by using a detection algorithm that fuses information gathered from the entire sensor network. The data facilitate good detection performance in part because the smoke was released from within the sensor network, creating a relatively strong contaminant presence. Therefore, in our analysis, we evaluate detection performance in the more difficult case when the contaminant signal is suppressed by a factor of 20. The suppressed data simulate sensors with reduced discrimination between contaminant and background—a more realistic model of common low-cost CB sensors.

Network False-Alarm Rate Reduction. In addition to supporting our CFD analysis, the recorded field-test data could also be used to study the false-alarm rates of high-density networks of CB sensors. Generally, it is thought that as more sensors are added to a network, the overall probability of false alarms of the whole network will increase. This is true, in the absence of further data processing, but even the application of simple rules can mitigate unnecessary false alarms while maintaining a constant probability of detection.

We ran experiments to examine how detection performance varies with the number of sensors included in the network. In order to measure performance, we set the alarm threshold low enough to achieve a specified benchmark for successful detections, then counted the false alarms triggered at that same threshold. We defined the false-alarm ratio as the number of false alarms divided by the total number of background data samples. Figure 11 shows the false-alarm ratio of the sensor network as a function of the number of available sensors. The curves for all three benchmarks exhibit the same general trend: we see a false-alarm reduction in the range of 2 to 3 orders of magnitude as the number of distributed sensors is increased from 1 to 23 (the maximum number of functioning particle counters for which data were collected).

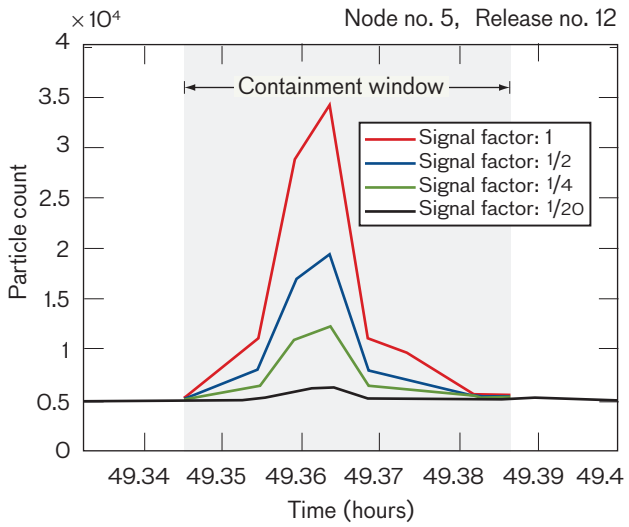


FIGURE 10. The graph shows the particle count time series for several different contaminant signal strengths, ranging from the original signal strength (factor 1) to low strength (factor 1/20). The number of particles due to background noise remains the same.

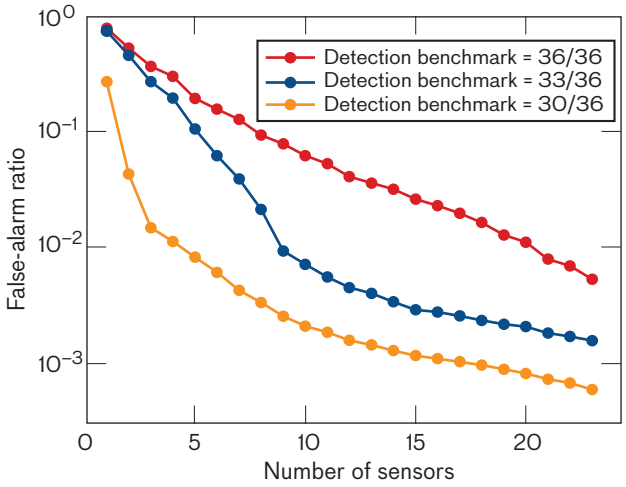


FIGURE 11. Increasing the number of sensors lowers the incidence of false alarms. Each curve plots the false-alarm ratios corresponding to a particular detection requirement for the 36 releases. The top curve plots the false-alarm ratio corresponding to a benchmark of perfect detection, while the lower two curves plot the false-alarm ratio corresponding to relaxed benchmarks of 33-out-of-36 and 30-out-of-36 successful detections, respectively. Since the placement of the sensors affects the experimental results, we averaged each curve across 200 trials, with randomly selected sensor locations for each trial.

Network Detection Algorithms

With a spatially distributed sensor network in place, one might enhance the network's overall detection performance by applying detection algorithms that operate across the entire network. These detection algorithms are able to identify spatial and temporal correlations that may help discriminate plume-like activity from random fractionations in the individual sensor readings. To explore the capabilities of network detection algorithms, we evaluated several algorithms on suppressed-signal field data. Such data present more of a detection challenge than do the original data. It is important to note that the field data exhibit several properties that require data preprocessing. The level of background noise drifts significantly throughout the time period of data collection because of varying particle concentrations in the ambient backgrounds. To deal with this noise drift, we apply two different forms of data normalization that extract more meaningful information from the particle counts, as described below:

Normalization method #1. At each time step during data collection, we take all particle counts reported anywhere in the network within a 30-minute window preceding that point in time. We compute the sample mean and standard deviation of the set of readings, then express the particle count at the current time step in units of standard deviations above or below the mean.

Normalization method #2. This method is similar to method #1, except that the set of preceding particle counts is specific to each sensor node. As a result, the readings are normalized by using sample means and standard deviations that are local in space as well as time. In addition to normalization, we also apply a simple anomaly detector to help screen out extreme particle counts that are due to sensor malfunction or short bursts of noise.

We considered several types of centralized contaminant detection algorithms. The simplest is the "max algorithm," which takes the maximum reading reported by any sensor node within a particular time frame as the indicator of contaminant presence (also called the "contaminant score"). The max algorithm, when applied to normalized sensor readings, performs well on most of the field data samples; however, when any one of the sensors experiences high levels of interferent noise, the contaminant score is inflated, causing false alarms. A better approach is to base the detection algorithm on a collection of network-wide features that better characterize the sensed phenomenon. Figure 12 shows an example of one such useful cue: spatial entropy, which measures the distribution of the particulate across the network.

We computed the following features for six instances, on the basis of readings reported during 16-second time frames: max readings of normalization methods #1 and #2, mean readings of normalization methods #1 and #2, spatial entropy of normalization method #2, and largest temporal gradient of normalization method #2. We used two different methods to derive a contaminant score from these six features. The first method builds Gaussian mixture models (GMM) [10] to characterize the distributions of both the contaminant data and the background data in the six-dimensional feature space. New data points are assigned a contaminant score by computing the ratio of the contaminant GMM value to the background GMM value. The second method computes the contaminant score of a data point by taking a weighted linear combination of its feature values. The weights of this linear discriminant (LD) are derived via iterative optimization so that contaminant data points yield high scores and background data points yield low scores [11]. The weight values, in

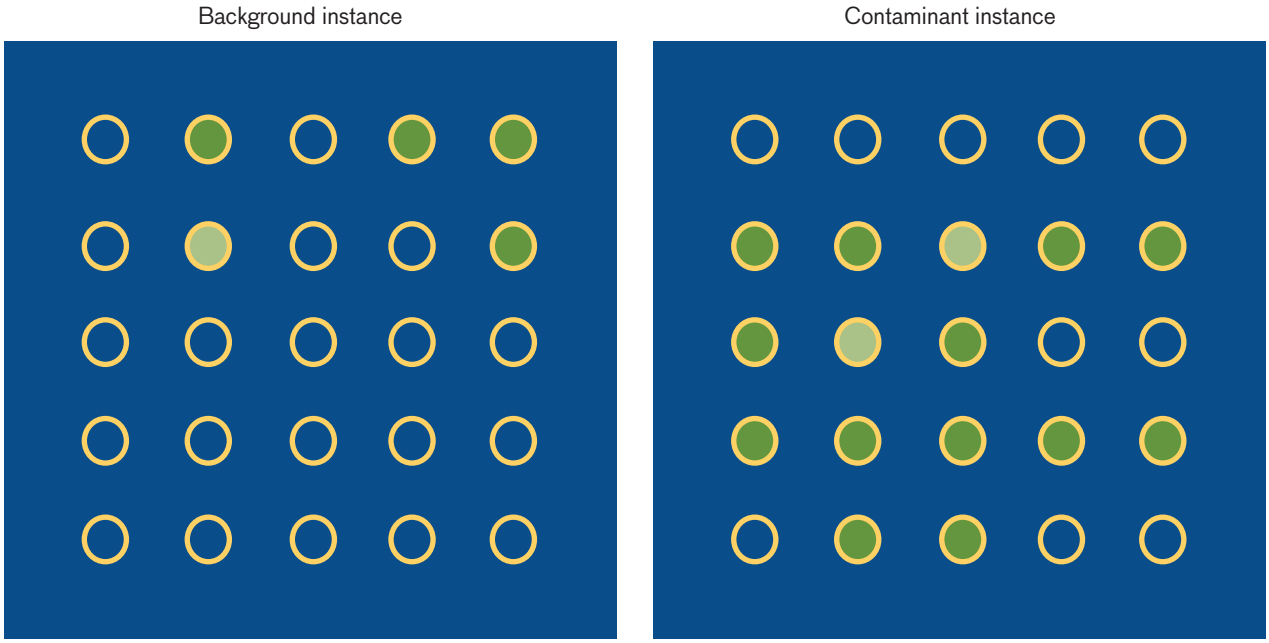


FIGURE 12. Particle count readings are shown, as reported by the 5-by-5 fielded sensor network within two different time frames. The green channel inside each circle is set proportional to the log of the normalized particle count for that sensor. **Left:** a noisy background instance with high maximum and low spatial entropy. **Right:** a contaminant release instance with lower maximum but higher spatial entropy.

combination with an alarm threshold, define a hyperplane that partitions the feature space into a contaminant region and a background region, as illustrated in Figure 13.

Table 1 summarizes detection results for the algorithms described above on adjusted field data, for which the contaminant signal strength has been suppressed by a factor of 20 (see Figure 10). The "baseline" algorithm, which employs no preprocessing or feature extraction, exhibits the poorest performance and clearly shows the need for some form of particle-count normalization. In addition, we get significant improvement by building classifiers in a feature space that better characterizes the joint information from multiple sensors, as demonstrated by the results of algorithms 3 and 4. We note that the detection methods in algorithms 3 and 4 each requires learning a small set of parameters; therefore, we evaluated these algorithms by averaging the results across two-fold cross-validation trials [12]. Figure 14 shows the algorithm performances graphically in the form of receiver-operator-characteristic curves. From these results, we conclude that merely establishing a distributed sensor network with good coverage does not guarantee significantly improved performance, especially when dealing with relatively weak contaminant signatures. As demonstrated in Table 1, the network's performance is critically dependent upon the algorithm used to fuse the information from the individual sensors.

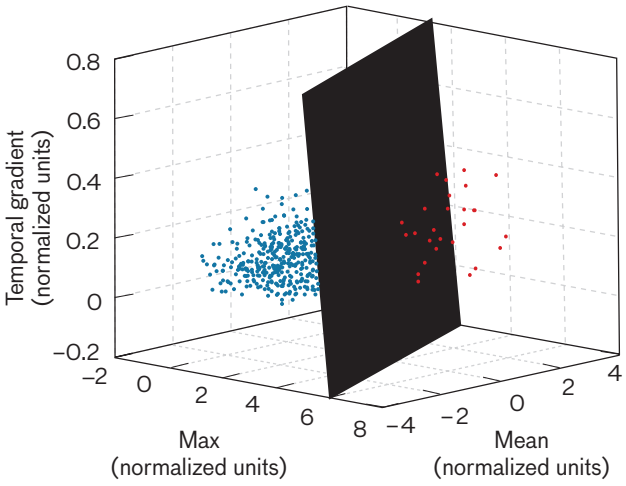


FIGURE 13. This visualization of the feature space used to characterize the sensor network data shows three out of six feature dimensions. A sample set of the field data is plotted, with each point representing the feature values measured within a single time frame. Red points indicate data collected during contaminant release periods of a field test; blue points indicate data collected during background periods. The algorithm learned to use a linear discriminant, shown as the plane in the figure, to separate the contaminant and background data.

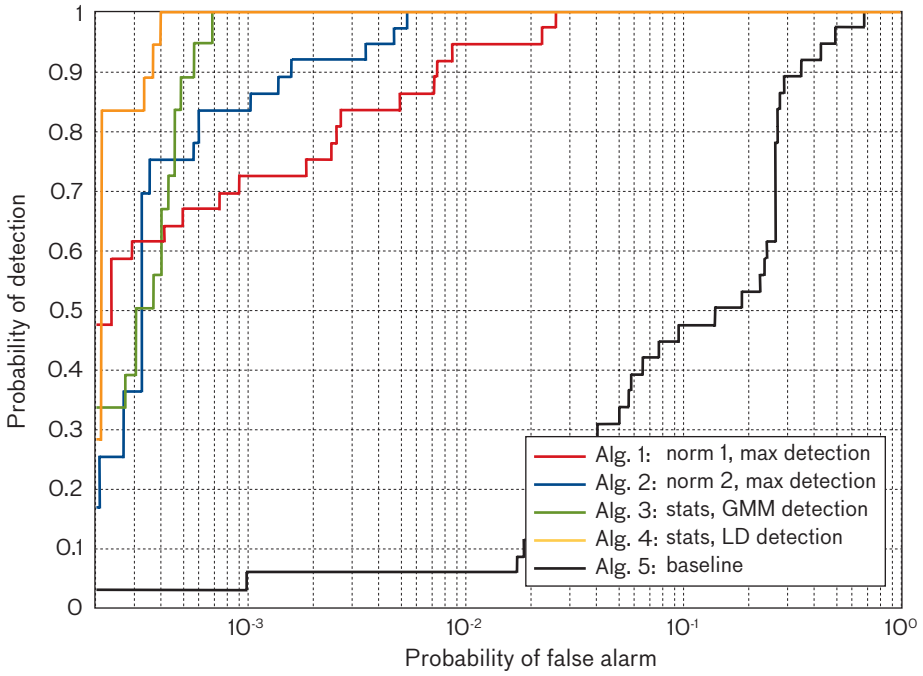


FIGURE 14. Receiver-operator-characteristic curves compare the performance of various detection algorithms. The baseline algorithm is displayed in black; the colored curves represent the algorithms performing detection by using the network's spatial and temporal features.

Communications Infrastructure

To pursue our vision of widely distributed, low-cost sensors for CB threats, we sought a communication infrastructure that was inexpensive and highly flexible. The possibilities offered by sensing based on "smart dust," more commonly known as "motes," were promising [13]. By employing motes within our sensing architecture, we were able to build a self-forming and adaptive wireless network while maintaining our low power consumption and low cost requirements.

Distributed Sensing and Wireless Communications. The vision of tiny cooperative sensors, so small as to be ubiquitous, is at least thirty years old (older if you include science fiction). However, it took more than twenty years of technology improvements before truly tiny, low-power sensors became an approachable target, and in order to have sensors cooperate, they must have communication elements as well. Advances in microelectromechanical systems (MEMS) and lower-cost manufacturing have brought many new devices and radios to a price and size/weight/power point appropriate for large-scale, long-term sensing deployments. It was for these reasons that we became interested in using an emerging communication platform, termed motes, for our data collection campaigns.

For our purposes, a mote is a computing and communication platform hosting one or more sensors [14, 15].

LABEL	PREPROCESSING METHOD	DETECTION METHOD	PROBABILITY OF FALSE ALARM
Baseline	None	Max algorithm	66.14%
Algorithm 1	Normalization method #1	Max algorithm	2.60%
Algorithm 2	Normalization method #2	Max algorithm	0.53%
Algorithm 3	Network-wide features	Gaussian mixture models	0.07%
Algorithm 4	Network-wide features	Linear discriminant	0.04%

TABLE 1. A comparison of the detection algorithm's probability of false alarm (for PD = 1) on field data with suppressed contaminant signal strength shows the need for some form of particle-count normalization. The performance of the baseline algorithm, which employs no preprocessing or feature extraction, is unacceptable. Significant improvement is obtained from classifiers that better characterize the joint information from multiple sensors, as shown by the results of algorithms 3 and 4.

Spatially distributed sensing applications require that some extra thought be given to data storage and communication. If each sensor's data can be stored locally and then collected after the deployment, then storage at each node will be adequate. However, if the data is to be observed in real time, or if data recorded at one location may be useful to observations made at another location, then the application requires a network capability.

Traditional wired networks do address this need and are often used for long-term fixed-site installations, but they have drawbacks. Because wiring increases the cost and effort of sensor emplacement, it is seldom practical for ad hoc deployments and never practical for mobile systems. A wireless networking scheme is more appropriate and is especially useful if it can adapt to a changing environment and changing sensor population without human intervention. The RF environment may change with the presence of interfering signals or reflective or absorbing materials, and hardware can fault or lose power. Mobile sensors, especially, may move in and out of range of their neighbors. When a single node is asked to transmit at a data rate that exceeds its capability, a bottleneck occurs and data may be dropped or delayed, thus degrading the performance of the detection algorithms. Therefore, designers of sensor networks must understand the capabilities and limits of the network architecture they employ. The following provides a brief overview of some typical mote network architectures.

Most sensing networks behave as local-area networks, with at least one base station or "gateway," which is a more provisioned node that provides a path for data exfiltration and command injection. Networks are typically described in terms of a connected graph, as shown in Figure 15. Simple scenarios, in which all nodes push data to (or receive commands or updates from) a base or master node, are common and easy to set up. When all nodes can communicate with the base station, the graph is often a simple "hub-spoke" configuration. Bottlenecks at the hub may be a concern, but the base station is often more fully provisioned (with a higher-power radio and deeper network stack) to prevent this. When not all nodes can maintain radio connectivity with the base, the structure becomes a "tree," wherein some nodes forward packets to and from outliers. In this case, bottlenecks may occur at the forwarding nodes, and link availability becomes an important concern during deployment, as

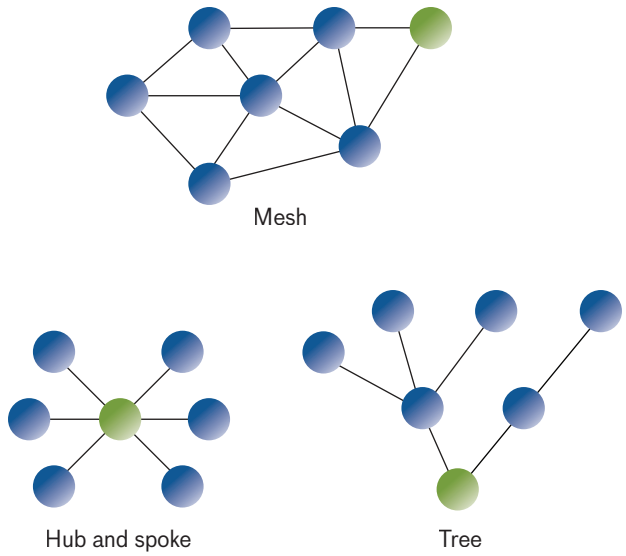


FIGURE 15. Connected graphs symbolize different network topologies: hub and spoke, tree, and mesh. The green nodes represent the base stations, which aggregate data.

nodes may need to be moved (increasing time) or added to the network (increasing cost, and sometimes sampling density). "Mesh" networks have more strongly connected graphs; in these networks, motes maintain information about several possible paths for data transmission and share sensor data with several of their neighbors. Reconfiguration happens more quickly when it is needed, and distributed algorithms requiring cooperation among sensors are more naturally supported. Finally, the hub-spoke, tree, and mesh topologies can be combined into "tiered" or "clustered" networks in which data is transmitted among local neighborhoods of nodes before being analyzed and passed up to a higher-level network or algorithm.

Because radio communication is a significant power consumer in the mote architecture, recent research has focused on ways to transmit data efficiently, such as synchronized duty-cycling of the radios, improved multi-hop routing protocols, and distributed detection algorithms (so that nodes don't need to pass as much raw data back to the base station). There is no silver-bullet architecture for distributed sensing: different deployments have different constraints and data-sharing requirements. We decided not to design a perfectly optimized network architecture for our system because we wanted to get a prototype sensor out in the field and learn what would work at a basic level.

Experimenting with Motes. In order to gain familiarity with motes and identify the capabilities of the technology, we conducted a series of experiments with various mote deployments. The general goal of these deployments was to determine the implications of using motes as a communication backbone for our distributed CB sensor networks.

During the experiments, we deployed two different commercially available platforms. The first was the Mica2 by Crossbow, Inc., which was equipped with a 433 MHz radio capable of 38.4 kbps and a maximum outdoor range of 300 m. The second was the M2135 by Dust Networks, which was equipped with a 2.4 GHz radio capable of 250 kbps with a maximum outdoor range of 400 m. When deployed, each mote platform automatically formed a mesh network, but the M2135 also used a proprietary time-synchronization protocol that significantly reduced the number of dropped data packets. During our experiments, we deployed test networks ranging in size from 20 to 90 mote nodes. We also used a network of twenty-five Mica2 motes to serve as the real-time communications network for our field tests. These experiments provided us with a good sense of the capabilities of modern low-power meshes. We learned that, in particular, power consumption, data latency, and network data convergence can have significant implications for high-density distributed CB sensing. Network data convergence strongly affects the overall detection performance of the sensor network. As the bandwidth of a low-power radio link is relatively small, data may be lost when too many nodes are trying to forward information to an aggregator node. This bottleneck reduces the data available to the network detection algorithm, and, in turn, reduces the algorithm's capability to suppress false alarms or identify threats.

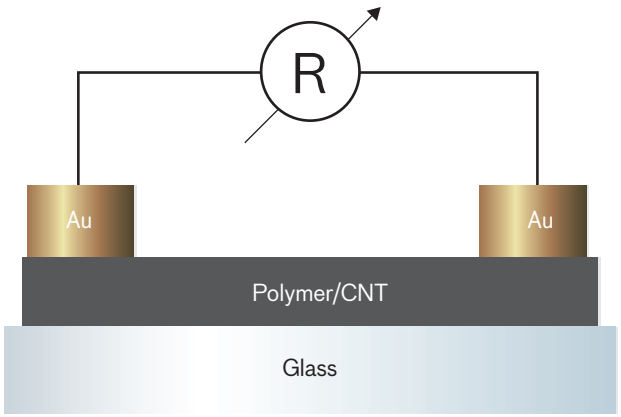


FIGURE 16. The polymer carbon nanotube (CNT) sensor is fabricated by patterning the hexafluoroisopropyl-substituted polythiophene (HFIP-PT) polymer on a glass substrate. Gold electrodes are bonded to the polymer/CNT layer to create electrical contacts. As polymer interaction with nerve agents causes a change in the polymer's electrical resistivity, we perform sensor readings by measuring the current flow through the polymer.

Inexpensive Sensor Technologies

The past decade has seen a surge in research and commercialization of devices capable of sensing chemical and biological warfare agents. Most devices have been developed for military applications, such as force protection and facility defense. However, these systems have frequently been adapted and marketed for homeland security and first-responder markets as well. Typically, military sensors are designed to issue rapid alerts in the presence of extremely low concentrations of specific agents of interest. In many instances, detection is limited to a small set of agents because of the requirement for high sensitivity while maintaining extremely low false-alarm rates. One specific defense application is the detection of a release of a threat agent at a military installation. This scenario entails the mobilization and positioning of complex sensors to achieve standoff sensing capabilities or, as standoff sensing is often impractical, the positioning of point sensors at very specific vulnerable locations. In both cases, the sensors require considerable amounts of manpower and training for accurate operation, and are extremely expensive. Even the best instruments are subject to false alarms from a variety of man-made or environmental sources, and are insensitive to trace amounts of truly toxic substances. Ideally, devices can be paired or grouped to decrease the false-alarm rate and increase the agent coverage; however, because of the high cost of traditional sensors, this approach is rare and sensors are often sparsely positioned.

The problem of limited sensor coverage would be overcome easily if these devices could be made inexpensively and with low power consumption. Ideally, they could be as widely deployed as commonly used smoke alarms. This approach would also permit the use of less sensitive sensors, as at least some of the sensors would be positioned at or near the agent source, where the threat would be sampled at much higher concentrations. Finally, if the threat could be tracked spatially as it moved across a grid of sensors, this correlated information could be

utilized to decrease false-alarm rates, identify the release location, provide early warning to those in the threat path, and determine the extent of the contaminated area.

Architecture-Compatible Sensor Requirements. A widely deployed grid of sensors requires a different set of sensor characteristics from that of traditional sensors. The specific capabilities of each sensor type to be incorporated into the network should ultimately be determined by the particular applications and operational environment. Field tests to monitor air flows provide us with valuable information that will lead to the understanding of sensor sensitivities and response times. In addition, the sensor power requirements and data processing must be suitable for long-term field use, and compatible with the grid's communication network. While few sensors have been designed specifically for field use under these guidelines, existing chemical and biological sensing constructs may address the requirements of low-to-moderate sensitivity, rapid response, discrimination at the threat class level, low maintenance, and low cost.

Candidate Biological Sensors. Traditional biological sensors that enable real-time detection of threats are typically based on the interrogation of individual particles using specific wavelengths of light to produce a scatter and fluorescence signature. These devices have been too expensive for widespread deployment in a distributed sensor network, however, mainly because of their need for a laser to provide the incident light. But newly developed light-emitting diodes have enabled the same type of detection performance at a much lower cost. Bioaerosol sensor developers strive for both rapid bioparticle discrimination and minimal maintenance during field use. Sensors offering additional discrimination among biological materials generally require the collection of samples followed by the addition of reagents for processing and identification. For the applications discussed here, this process is not acceptable because of the lengthened processing time (several minutes under ideal conditions), the additional reagents and implied maintenance, and the cost of the sample collection and processing hardware.

Candidate Chemical Sensors. The technology that holds the most promise for chemical detection is polymer-based sensing. Polymer sensors are inexpensive to fabricate and the polymer can be designed to specifically bind single compounds or classes of chemicals. The sensor readout from a polymer platform can use a number

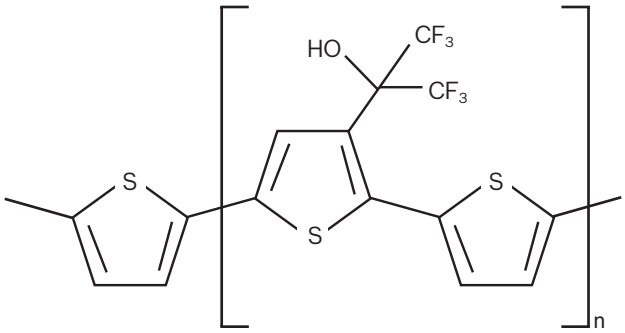


FIGURE 17. HFIP-PT, a chemoresistive-conducting polymer, functions as a dispersing medium for carbon nanotubes. Lincoln Laboratory is incorporating this film into distributed sensor networks.

of different transduction schemes, including surface acoustic wave, microcantilevers, and fluorescence, to name a few. The simplest candidate is the chemiresistor: a chemical sensor that exploits a measurable change in resistance in response to an analyte binding event, in this case chemical warfare agents. The simplicity and flexibility of this transduction mechanism conveys reasonable detection sensitivity and maintains the low power and cost requirements, so it suits a large deployment. One depiction of this sensing scheme is given in Figure 16, which shows the general layout of electrodes and bridging material. Many materials (e.g., metal oxides, organic semiconductors) are sensitive to analyte binding events; the recent focus of many researchers on developing conductive polymer/carbon nanotube (CNT) films has produced promising materials.

The synthetic challenge for designing effective sensors using chemiresistor technology is twofold. First, the designer must choose polymers that are responsive to the analytes of interest but not responsive to typical environmental changes that cannot be controlled in a field setting. Second, the designer must maintain control of the signal conduction pathways so that a measurable signal can be observed with appropriate concentrations of analyte. A dispersion of single-walled CNTs in a conductive polymer film enables optimization of both of these requirements.

Inexpensive Sensor Development. In an ongoing collaboration with Prof. Timothy Swager at MIT, Lincoln Laboratory acquired a chemoresistive conducting polymer/CNT film that used a high-molecular-weight hexafluoroisopropanol-substituted polythiophene (HFIP-PT) as a dispersing medium for single-walled CNTs, see Figure

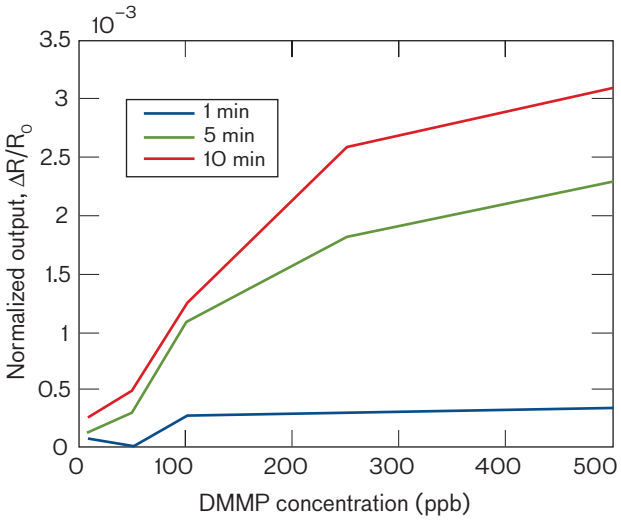


FIGURE 18. Polymer sensor responds to increasing concentrations of dimethyl methylphosphonate (DMMP). We exposed the polymer sensor to different levels of DMMP concentration for three different exposure times. These results show that the sensor is capable of responding to small analyte concentrations within a few minutes.

17 [16, 17]. This polymer/CNT film is being incorporated into a fieldable sensor platform at Lincoln Laboratory for use in distributed sensor networks. Solution-casting this material created a noncontinuous, yet conductive, film wherein CNTs defined conduction pathways and the polythiophene provided conductive junctions. The initial choice of polythiophene was based on MIT's extensive experience with the polymer, in addition to widespread knowledge of the polymer's properties. The polythiophene backbone was then functionalized with HFIP to produce selective responses to the organophosphate-based nerve gases known as G-agents. HFIP was chosen because of its hydrogen bonding with phosphonate esters common to the G-agents and because of its minimal sensitivity to moisture levels in the air (a common background interferent that plagues many polymer-based sensors) and other common interferent classes of chemicals (e.g., aromatic hydrocarbons, alkanes, halogenated compounds, and alcohols). Preliminary examination of the polymer/CNT film was completed at MIT with the sarin simulant dimethyl methylphosphonate (DMMP) and continued through the use of a testbed developed for controlled testing against relevant environmental factors (temperature and humidity) and potential interferents. Figure 18 shows the polymer sensor's response to varied concentrations of DMMP with increasing exposure times. Approximately 50 ppb of DMMP is readily detected when the polymer is exposed to the vapor for 5 minutes. Sensor testing under environmental conditions and to examine sensor stability is currently under way.

To develop sensors appropriate for large-scale deployment, it is critical to establish a simple, reproducible fabrication process and incorporate sensor component hardware that does not inflate the overall sensor cost. Thus, we have refined the sensor packaging hardware and electronics in order to generate a robust fieldable sensor with optimal response characteristics.

Next Steps

High-density distributed sensing for CB defense is still in its infancy. Our experimentation and analysis efforts displayed the effectiveness of a distributed sensing framework and explored the initial development of network detection algorithms, low-power wireless communications, and low-cost sensor technologies. To advance each of these fields, we are currently exploring the following topics:

- The development of advanced detection capabilities such as the tracking, mapping, and prediction of CB plumes
- Exploring the effect of network data convergence and data loss on the sensor network's overall detection performance
- The continued development of low-cost chemical sensors

The information retrieved from spatially distributed CB sensors enables operators to determine not only the presence of a threat, but also how the threat propagates through space. With proper analysis, we can use this spatial information to map contaminated areas, track the plume's current location, and predict future hazard regions. These advanced detection capabilities will not only significantly enhance operators' situational awareness, but may also assist efforts of source attribution by enhancing other capabilities, such as plume source localization. To develop these capabilities, we are currently investigating the use of wind-induced correlations across sensor readings to derive tracking, mapping, and prediction capabilities.

As we discovered in our experiments with low-power mesh networking radios, the transmission of sensor data throughout the sensor network can affect the network's

overall detection performance. In-network processing or network intermittency may reduce the amount of sensor data that reaches a central data processor. When this occurs, the network detection algorithm operates with less information and may reduce the overall detection performance. We are currently designing simulations to quantify the implications of data loss on detection performance. We are also exploring the trade-offs involved with performing in-network computation versus running the detection algorithms from a centralized computer.

Prior CB sensor development has focused on building the next-generation high-fidelity sensor, but the advent of high-density distributed sensing enables the practical and effective use of low-cost sensors. Unlike the high-fidelity CB sensor class, this domain remains underdeveloped, with only a few representative sensors to meet a broad need. To address the requirements of other applications, we are pursuing the development of another low-cost chemical sensor, in addition to the discussed polymer sensor.

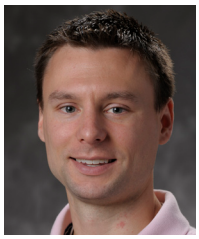
Acknowledgments

The authors thank Rachel Achituv, Derek Young, Mindy Reynolds, and Matthew Peters for their technical contributions to this work. We thank Prof. Timothy Swager of MIT for the development and production of his conductive polymers. We also thank Bernadette Johnson for her support and guidance throughout this effort. ■

REFERENCES

1. D. Cousins and S.D. Campbell, "Protecting Buildings against Airborne Contamination," *Linc. Lab. J.*, vol. 17, no. 1, 2007, pp. 131–151.
2. J. Carrano, "Chemical and Biological Sensor Standards Study," Technical Report, Defense Advanced Research Projects Agency, Arlington, Va., Aug. 2007.
3. C.J. Nappo, "Turbulence and Dispersion Parameters Derived from Smoke-Plume Photoanalysis," *Atmos. Environ.*, vol. 18, no. 2, 1984, pp. 299–306.
4. S. Bradley, T. Mazzola, R. Ross, D. Srinivasa, and R. Fry, "Initial verification and validation of HPAC 1.3," Technical Report, Defense Special Weapons Agency, Alexandria, Va., Nov. 1997, ADB231209. Export Control Distribution authorized to U.S. government agencies and their contractors.
5. J.C. Chang, P. Franzese, K. Chayantrakom, and S.R. Hanna, "Evaluations of Calpuff, HPAC, and VLSTRACK with Two Mesoscale Field Datasets," *J. Appl. Meteorol.*, vol. 42, no. 4, 2003, pp. 453–466.
6. W.J. Coirier and S. Kim, "CFD Modeling for Urban Area Contaminant Transport and Dispersion: Model Description and Data Requirements," *JP2.11*, 86th AMS Annual Meeting, Sixth Symposium on the Urban Environment and Forum on Managing Physical and Natural Resources, American Meteorological Society, 2006.
7. W.J. Coirier, S. Kim, S.C. Ericson, and S. Marella, "Calibration and Use of Site-Specific Urban Weather Observations Data Using Microscale Modeling," 87th AMS Annual Meeting, 14th Symposium on Meteorological Observation and Instrumentation, American Meteorological Society, Jan. 2007.
8. K.J. Allwine, M.J. Leach, L.W. Stockham, J.S. Shinn, R.P. Hosker, J.F. Bowers, and J.C. Pace, "Overview of Joint Urban 2003—An Atmospheric Dispersion Study in Oklahoma City." *Bull. Am. Meteorol. Soc.*, Combined Preprints: 84th American Meteorological Society (AMS) Annual Meeting, 2004, pp. 745–753.
9. A.T. Tu, "Basic Information on Nerve Gas and the Use of Sarin by Aum Shinrikyo: Plenary Lecture at the Biological Mass Spectrometry Conference, Seto, Aichi, Jpn., July 3–6, 1995," *J. Mass Spectro. Soc. Japan*, vol. 44, no. 3, 1996, pp. 293–320.
10. D. Titterington, A.F.M. Smith, and U.E. Makov, *Statistical Analysis of Finite Mixture Distributions,* Chichester, New York: Wiley, 1985.
11. G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition,* Hoboken, N.J.: Wiley-InterScience, 2004.
12. R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," *Proc. 14th Int. Joint Conf. Artific. Intell., IJCAI-95*, vol. 2, no. 12, 1995, pp. 1137–1143.
13. B. Warneke, M. Last, B. Liebowitz, and K.S.J. Pister, "Smart Dust: Communicating with a Cubic-Millimeter Computer," *Computer*, vol. 34, no. 1. 2001, pp. 44–51.
14. D. Gay, P. Levis, R. von Behren, M. Welsh, E. Brewer, and D. Culler, "The nesC Language: A Holistic Approach to Networked Embedded Systems," *Proc. ACM SIGPLAN Conf. Prog. Lang. Design and Impl. (PLDI)*, 2003, pp. 1–11.
15. P. Levis, *TinyOS Programming*, web published, 2006. http://www.tinyos.net/tinyos-2.x/doc/pdf/tinyos-programming.pdf.
16. F. Wang, H. Gu, and T.M. Swager, "Carbon Nanotube/Polythiophene Chemiresistive Sensors for Chemical Warfare Agents," *J. Am. Chem. Soc.*, vol. 130, no. 16, 2008, pp. 5392–5393.
17. K. Sugiyasu and T.M. Swager, "Conducting-Polymer-based Chemical Sensors: Transduction Mechanisms," *Bull. Chem. Soc. Jpn.*, vol. 80, no. 11, 2007, pp. 2074–2083.

ABOUT THE AUTHORS



Adam Norige is an associate staff member in the Biodefense Systems Group, where his work includes the development of systems for the detection of biological and chemical defense as well as the development of strategic decision-making technologies for disaster response. He received bachelor's and master's degrees in biomedical engineering from Worcester Polytechnic Institute, where he researched multilayered oxygen tension maps of the retina.

Jason Thornton is a staff member in the Biodefense Systems Group, where he works on algorithms for pattern recognition, sensor fusion, and decision support. He holds a doctorate in electrical and computer engineering from Carnegie Mellon University, and a bachelor's degree in information and computer science from the University of California, Irvine.

Curran Schiefelbein is an associate staff member in the Biodefense Systems Group, where she develops distributed sensor networks for advanced warning of biological or chemical attacks. She has bachelor's and master's degrees in computer science from Brown University.

Christina Rudzinski is a staff member in the Biodefense Systems Group. She is developing systems to detect chemical, biological, and explosive weapons in maritime shipping containers and designing novel reagents to enhance biological agent detection and identification from complex clinical and environmental samples. She has a doctorate in inorganic chemistry from MIT, where she designed and studied molecules capable of highly specific and sensitive fluorescence-based detection.